

SEQUENCE LISTING

(1) GENERAL INFORMATION:

5

(i) APPLICANT: Griffith, Irwin J.
Pollock, Joanne, Bond Julian

10.

(ii) TITLE OF INVENTION: Allergenic Proteins And Peptides From
Japanese Cedar Pollen

(iii) NUMBER OF SEQUENCES: 25

(iv) CORRESPONDENCE ADDRESS:

15

(A) ADDRESSEE: ImmuLogic Pharmaceutical Corporation
(B) STREET: One Kendall Square, Building 600
(C) CITY: Boston
(D) STATE: MA
(E) COUNTRY: USA
20 (F) ZIP: 02139

(v) COMPUTER READABLE FORM:

25

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

35

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Stacey L. Channing
(B) REGISTRATION NUMBER: 31,095
(C) REFERENCE/DOCKET NUMBER: IPC-025CC

40

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (617) 494-0060
(B) TELEFAX: (617) 494-4964

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1337 base pairs
- (B) TYPE: nucleic acid
- 5 (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

10 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Cryptomeria japonica

(ix) FEATURE:

- (A) NAME/KEY: CDS
- 15 (B) LOCATION: 66..1187

(ix) FEATURE:

- (A) NAME/KEY: mat_peptide
- (B) LOCATION: 129..1187

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

25	AGTCAATCTG CTCATAATCA TAGCATAGCC GTATAGAAAG AAATTCTACA CTCTGCTACC	60
	AAAAA ATG GAT TCC CCT TGC TTA GTA GCA TTA CTG GTT TTC TCT TTT	107
	Met Asp Ser Pro Cys Leu Val Ala Leu Leu Val Phe Ser Phe	
	-21 -20 -15 -10	
30	GTA ATT GGA TCT TGC TTT TCT GAT AAT CCC ATA GAC AGC TGC TGG AGA	155
	Val Ile Gly Ser Cys Phe Ser Asp Asn Pro Ile Asp Ser Cys Trp Arg	
	-5 1 5	
35	GGA GAC TCA AAC TGG GCC CAA AAT AGA ATG AAG CTC GCA GAT TGT GCA	203
	Gly Asp Ser Asn Trp Ala Gln Asn Arg Met Lys Leu Ala Asp Cys Ala	
	10 15 20 25	
40	GTG GGC TTC GGA AGC TCC ACC ATG GGA GGC AAG GGA GAT CTT TAT	251
	Val Gly Phe Gly Ser Ser Thr Met Gly Gly Lys Gly Asp Leu Tyr	
	30 35 40	
45	ACG GTC ACG AAC TCA GAT GAC GAC CCT GTG AAT CCT GCA CCA GGA ACT	299
	Thr Val Thr Asn Ser Asp Asp Pro Val Asn Pro Ala Pro Gly Thr	
	45 50 55	
50	CTG CGC TAT GGA GCA ACC CGA GAT AGG CCC CTG TGG ATA ATT TTC AGT	347
	Leu Arg Tyr Gly Ala Thr Arg Asp Arg Pro Leu Trp Ile Ile Phe Ser	
	60 65 70	
	GGG AAT ATG AAT ATA AAG CTC AAA ATG CCT ATG TAC ATT GCT GGG TAT	395
	Gly Asn Met Asn Ile Lys Leu Lys Met Pro Met Tyr Ile Ala Gly Tyr	

	75	80	85	
5	AAG ACT TTT GAT GGC AGG GGA GCA CAA GTT TAT ATT GGC AAT GGC GGT Lys Thr Phe Asp Gly Arg Gly Ala Gln Val Tyr Ile Gly Asn Gly Gly 90 95 100 105			443
10	CCC TGT GTG TTT ATC AAG AGA GTT AGC AAT GTT ATC ATA CAC GGT TTG Pro Cys Val Phe Ile Lys Arg Val Ser Asn Val Ile Ile His Gly Leu 110 115 120			491
15	TAT CTG TAC GGC TGT AGT ACT AGT GTT TTG GGG AAT GTT TTG ATA AAC Tyr Leu Tyr Gly Cys Ser Thr Ser Val Leu Gly Asn Val Leu Ile Asn 125 130 135			539
20	GAG AGT TTT GGG GTG GAG CCT GTT CAT CCT CAG GAT GGC GAT GCT CTT Glu Ser Phe Gly Val Glu Pro Val His Pro Gln Asp Gly Asp Ala Leu 140 145 150			587
25	ACT CTG CGC ACT GCT ACA AAT ATT TGG ATT GAT CAT AAT TCT TTC TCC Thr Leu Arg Thr Ala Thr Asn Ile Trp Ile Asp His Asn Ser Phe Ser 155 160 165			635
30	AAT TCT TCT GAT GGT CTG GTC GAT GTC ACT CTT ACT TCG ACT GGA GTT Asn Ser Ser Asp Gly Leu Val Asp Val Thr Leu Thr Ser Thr Gly Val 170 175 180 185			683
35	ACT ATT TCA AAC AAT CTT TTT TTC AAC CAT CAT AAA GTG ATG TTG TTA Thr Ile Ser Asn Asn Leu Phe Phe Asn His His Lys Val Met Leu Leu 190 195 200			731
40	GGG CAT GAT GAT GCA TAT AGT GAT GAC AAA TCC ATG AAG GTG ACA GTG Gly His Asp Asp Ala Tyr Ser Asp Asp Lys Ser Met Lys Val Thr Val 205 210 215			779
45	GCG TTC AAT CAA TTT GGA CCT AAC TGT GGA CAA AGA ATG CCC AGG GCA Ala Phe Asn Gln Phe Gly Pro Asn Cys Gly Gln Arg Met Pro Arg Ala 220 225 230			827
50	CGA TAT GGA CTT GTA CAT GTT GCA AAC AAT AAT TAT GAC CCA TGG ACT Arg Tyr Gly Leu Val His Val Ala Asn Asn Tyr Asp Pro Trp Thr 235 240 245			875
55	ATA TAT GCA ATT GGT GGG AGT TCA AAT CCA ACC ATT CTA AGT GAA GGG Ile Tyr Ala Ile Gly Gly Ser Ser Asn Pro Thr Ile Leu Ser Glu Gly 250 255 260 265			923
60	AAT AGT TTC ACT GCA CCA AAT GAG AGC TAC AAG AAG CAA GTA ACC ATA Asn Ser Phe Thr Ala Pro Asn Glu Ser Tyr Lys Lys Gln Val Thr Ile 270 275 280			971
65	CGT ATT GGA TGC AAA ACA TCA TCA TCT TGT TCA AAT TGG GTG TGG CAA Arg Ile Gly Cys Lys Thr Ser Ser Cys Ser Asn Trp Val Trp Gln 285 290 295			1019
70	TCT ACA CAA GAT GTT TTT TAT AAT GGA GCT TAT TTT GTA TCA TCA GGG Ser Thr Gln Asp Val Phe Tyr Asn Gly Ala Tyr Phe Val Ser Ser Gly 300 305 310			1067
75	AAA TAT GAA GGG GGT AAT ATA TAC ACA AAG AAA GAA GCT TTC AAT GTT			1115

(2) INFORMATION FOR SEQ ID NO:2:

- 20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 374 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

25 (ii) MOLECULE TYPE: protein

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Asp	Ser	Pro	Cys	Leu	Val	Ala	Leu	Leu	Val	Phe	Ser	Phe	Val	Ile
-21	-20						-15							-10	
Gly Ser Cys Phe Ser Asp Asn Pro Ile Asp Ser Cys Trp Arg Gly Asp															
-5 1 5 10															
Ser Asn Trp Ala Gln Asn Arg Met Lys Leu Ala Asp Cys Ala Val Gly															
15 20 25															
Phe Gly Ser Ser Thr Met Gly Gly Lys Gly Asp Leu Tyr Thr Val															
30 35 40															
Thr Asn Ser Asp Asp Asp Pro Val Asn Pro Ala Pro Gly Thr Leu Arg															
45 50 55															
Tyr Gly Ala Thr Arg Asp Arg Pro Leu Trp Ile Ile Phe Ser Gly Asn															
60 65 70 75															
Met Asn Ile Lys Leu Lys Met Pro Met Tyr Ile Ala Gly Tyr Lys Thr															
80 85 90															
Phe Asp Gly Arg Gly Ala Gln Val Tyr Ile Gly Asn Gly Gly Pro Cys															
95 100 105															
Val Phe Ile Lys Arg Val Ser Asn Val Ile Ile His Gly Leu Tyr Leu															
110 115 120															
Tyr Gly Cys Ser Thr Ser Val Leu Gly Asn Val Leu Ile Asn Glu Ser															
125 130 135															
Phe Gly Val Glu Pro Val His Pro Gln Asp Gly Asp Ala Leu Thr Leu															

140 145 150 155
Arg Thr Ala Thr Asn Ile Trp Ile Asp His Asn Ser Phe Ser Asn Ser
160 165 170
5 Ser Asp Gly Leu Val Asp Val Thr Leu Thr Ser Thr Gly Val Thr Ile
175 180 185
10 Ser Asn Asn Leu Phe Phe Asn His His Lys Val Met Leu Leu Gly His
190 195 200
Asp Asp Ala Tyr Ser Asp Asp Lys Ser Met Lys Val Thr Val Ala Phe
205 210 215
15 Asn Gln Phe Gly Pro Asn Cys Gly Gln Arg Met Pro Arg Ala Arg Tyr
220 225 230 235
Gly Leu Val His Val Ala Asn Asn Asn Tyr Asp Pro Trp Thr Ile Tyr
240 245 250
20 Ala Ile Gly Gly Ser Ser Asn Pro Thr Ile Leu Ser Glu Gly Asn Ser
255 260 265
Phe Thr Ala Pro Asn Glu Ser Tyr Lys Lys Gln Val Thr Ile Arg Ile
270 275 280
25 Gly Cys Lys Thr Ser Ser Cys Ser Asn Trp Val Trp Gln Ser Thr
285 290 295
30 Gln Asp Val Phe Tyr Asn Gly Ala Tyr Phe Val Ser Ser Gly Lys Tyr
300 305 310 315
Glu Gly Gly Asn Ile Tyr Thr Lys Lys Glu Ala Phe Asn Val Glu Asn
320 325 330
35 Gly Asn Ala Thr Pro Gln Leu Thr Lys Asn Ala Gly Val Leu Thr Cys
335 340 345
40 Ser Leu Ser Lys Arg Cys
350
45 (2) INFORMATION FOR SEQ ID NO:3:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 17 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
50
55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:
GAYAAYCCNA THGAYWS
17

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
5 (B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GGGAATTCAA YTGGGCNCAR AAYSG

25

15 (2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
20 (B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

25 (ix) FEATURE:

- (A) NAME/KEY: modified_base
(B) LOCATION: 15
(D) OTHER INFORMATION: /mod_base= i

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CTGCAGCCRT TYTCNACRTT RAA

23

35 (2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
40 (B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

45 (ix) FEATURE:

- (A) NAME/KEY: modified_base
(B) LOCATION: 6
(D) OTHER INFORMATION: /mod_base= i

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TTCATNCKRT TYTGNGCCCA

20

5

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

10

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CCTGCAGCKR TTYTGNGCCC AARTT

25

20

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

ATGGATTCCC CTTGCTTA

18

35

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

40

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GGGAATTCGA TAATCCCAT A GACAGC

26

50

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
5 (A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

ATGCCTATGT ACATTGC

17

15

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
20 (A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GCAATGTACA TAGGCAT

17

30

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
35 (A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TCCAATTCTT CTGATGGT

18

45

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
50 (A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TTTTGTCAAT TGAGGAGT

18

10

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CCTGCAGAAG CTTCATCAAC AACGTTTAGA

30

25

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

TAGCAACTCC AGTCGAAGT

19

40

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

TAGCTCTCAT TTGGTGC

17

5

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

10

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

TATGCAATTG GTGGGAGT

18

20

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

25

(ii) MOLECULE TYPE: peptide

30

(v) FRAGMENT TYPE: N-terminal

35

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Cryptomeria japonica

40

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 7
- (D) OTHER INFORMATION: /note= "the amino acid at

position

7 is Ser, Cys, Thr, or His"

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

50

Asp Asn Pro Ile Asp Ser Xaa Trp Arg Gly Asp Ser Asn Trp
Ala Gln

1

5

10

15

Asn Arg Met Lys

20

(2) INFORMATION FOR SEQ ID NO:19:

5 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

15 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Cryptomeria japonica

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

20 Glu Ala Phe Asn Val Glu Asn Gly Asn Ala Thr Pro Gln Leu

Thr Lys

1

5

10

15

25 (2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GGGTCTAGAG GTACCGTCCG ATCGATCATT

30

40 (2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GGGTCTAGAG GTACCGTCCG
20

(2) INFORMATION FOR SEQ ID NO:22:

5

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
10 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

15

AATGATCGAT GCT
13

(2) INFORMATION FOR SEQ ID NO:23:

20

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

30

GGAATTCTCT AGACTGCAGG T
21

(2) INFORMATION FOR SEQ ID NO:24:

35

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
40 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

45

GGAATTCTCT AGACTGCAGG TTTTTTTTT TTTTT 35

(2) INFORMATION FOR SEQ ID NO:25:

50

- (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: N-terminal

10 (vi) ORIGINAL SOURCE:

(A) ORGANISM: Juniperus sabinaoides

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

15 Asp Asn Pro Ile Asp
1 5